Modeling the Effects of Uncertainty on the National Airspace System

Erin C. DeCarlo[1] and Barron J. Bichon[2]

[1,2]*Southwest Research Institute, San Antonio, TX, 78238, U.S.A*
erin.decarlo@swri.org
barron.bichon@swri.org

ABSTRACT

This paper presents the first steps toward managing uncertainty and assessing risk within the national airspace system (NAS) by investigating the impact of uncertainty on "flight plan flexibility" (FPF) – a proposed quantitative measure of an aircraft's ability to adapt its flight plan due to improbable events. First, an air traffic scenario derived from national flight plan data is simulated with an open source BlueSky air traffic control analysis centered on a busy airport. Next, state-space diagrams derived from the aircraft state parameters (i.e., speeds, altitudes, headings), spatial proximities, and surveillance signals are used to construct the FPF metric. Finally, a probabilistic analysis is used to propagate uncertainty in the aircraft positions through BlueSky to observe the resulting uncertainty in FPF through time. Future work will aggregate individual aircraft safety measures and additional metrics into a single system-wide indicator, transitioning from BlueSky to a gate-to-gate simulation for prognostics, and deriving probabilistic models of epistemic and aleatory sources of uncertainty in the NAS from available data.

1. INTRODUCTION

The control and management of the national airspace system (NAS) is projected to transition to advanced surveillance systems that will enable continuous system-wide safety monitoring and decision-making by 2025. One challenge is systematically integrating information from heterogeneous data sources (i.e., real-time and historical, quantitative and linguistic) into a comprehensive gate-to-gate NAS simulation for prognostics. Furthermore, a large amount of uncertainty surrounds incoming information from various streams of data (e.g., aeronautical instruments, voice communications, weather forecasts) in addition to inherent variability in manufactured aircraft components and flight environments.

To address these challenges, a 5-year project sponsored by the NASA University Leadership Initiative (ULI) and led by Arizona State University has been launched with the goal of advancing and integrating information fusion methodologies for real-time system-wide safety assurance of the NAS.

The primary goal of this project is to create a framework with which to fuse the multitude of existing and NextGen aviation data sources for prognostics within the NAS. These data sources currently include federally maintained databases such as the System-Wide Information Management (SWIM) and the Aviation Safety Reporting System data. In addition, data sources include the Base of Aircraft Data (BADA) of aircraft performance models developed by EUROCONTROL in cooperation with aircraft manufacturers and the real-time Automatic Dependent Surveillance-Broadcast (ADS-B) reports of aircraft positions and velocities (lateral and vertical). These information sources, among others, are to inform a gate-to-gate simulation framework for measuring the safety and risk in the NAS. This simulation will consider nominal and off-nominal procedural communications between air traffic control (ATC) and pilots, their respective performance and decision-making abilities under fatigued and non-fatigued conditions, as well as the overall risk contained in the air space due to the air traffic configuration and health of flight equipment (air frame, engines, sensors, etc).

Southwest Research Institute (SwRI) is the principal investigative organization on the "Uncertainty Management and Risk Assessment" task, which focuses on identifying and modeling all sources of uncertainty and variation within the NAS. Because reduced safety margins are expected with both increased demand on the NAS and enabling technologies such as ADS-B, a probabilistic modeling framework for large complex networks such as the NAS that is vulnerable to rare but potentially catastrophic events is needed. The work presented here represents the first steps toward creating the infrastructure for probabilistic analysis of the NAS and demonstrates the integration of the available models and data together for uncertainty quantification and risk assessment.

The gate-to-gate simulation framework is one of the long-term deliverables of the project, so this initial effort instead uses the open source ATC simulator BlueSky developed by researchers at Delft University of Technology (Hoekstra and Ellerbroek, 2016) for modeling air traffic, air traffic control commands, and conflict detection and resolution (CD&R). The CD&R strategy employed in this work is a state-space diagram (SSD)-based algorithm informed by an aircraft's forbidden velocities (to avoid conflict and a loss of separation (LoS)) and reachable velocities (from the aircraft performance limits) (Balasooriyan, 2017). The SSDs for each aircraft are available within BlueSky and serve as the basis of the "flight plan flexibility" (FPF) metric developed to quantify the vulnerability of the NAS.

There are categorical metrics for the position and velocity accuracy reported by ADS-B communications where each category corresponds to a level of uncertainty about the measurement. Uncertainty in the position and velocity of the aircraft leads to uncertainty about the computed FPF indicator. To study the effect of ADS-B uncertainty on the FPF metric, uncertainty in the aircraft position is propagated through BlueSky using NESSUS®, a SwRI-developed probabilistic analysis software that connects to executable simulations for uncertainty quantification, sensitivity analyses, and probabilistic reliability analysis of engineering systems.

The paper proceeds with Section 2 deriving the air traffic scenario from the SWIM database to simulate in BlueSky. Next, Section 3 constructs the FPF metric for NAS flexibility and vulnerability. Uncertainty in position due to ADS-B precision is introduced in Section 4 and propagated through a BlueSky air traffic scenario undergoing CD&R using NESSUS. Section 5 shows the results of the uncertainty analysis and Section 6 outlines research directions of interest going forward.

2. AIR TRAFFIC SCENARIO FROM SWIM DATA

The System-Wide Information Management (SWIM) system is administered by the Federal Aviation Administration (FAA) and provides a continuous feed of instantaneous flight plan messages over a specific time duration. SWIM flight plan messages were collected for increasing durations of time. Ultimately, a 24-hour dataset was necessary because the International Civil Aviation Organization (ICAO) aircraft model types are only reported once per flight plan. The aircraft type is needed as an input to the BlueSky air traffic simulation in order utilize BADA aircraft performance files. It is important to note that the SWIM messages are not produced in real-time; they correspond to the flights almost exactly a week prior (within seconds).

With the goal of simulating the traffic around a major airport in BlueSky, the data were filtered by those flight plans that entered the airspace within 100km (54 nautical miles (NM)) of Hartsfield-Jackson Atlanta International Airport (ATL), which is classified as the world's busiest airport and operates as the primary Delta Airlines (DAL) hub and a major hub for many others. Figure 1 shows a snapshot of the BlueSky simulation of the air traffic around ATL with the initial location of each aircraft and their speed, heading, and altitude defined by the SWIM dataset. Note that since an SSD-based CD&R strategy is used in this simulation, the air traffic positions are no longer a reflection of what actually occurred after the initial condition, however, the aircraft destinations from the SWIM data are guiding their simulated routes.

An additional consideration when looking at flight plans around ATL are the landing and takeoff directions of each aircraft. These directions are defined by the runway designations, which are also needed to study runway incursion in future work. Figure 2 depicts BlueSky's layout of ATL airport with the parallel runways (from top to bottom) being 08L↔26R, 08R↔26L, 09L↔27R, 09R↔27L, and
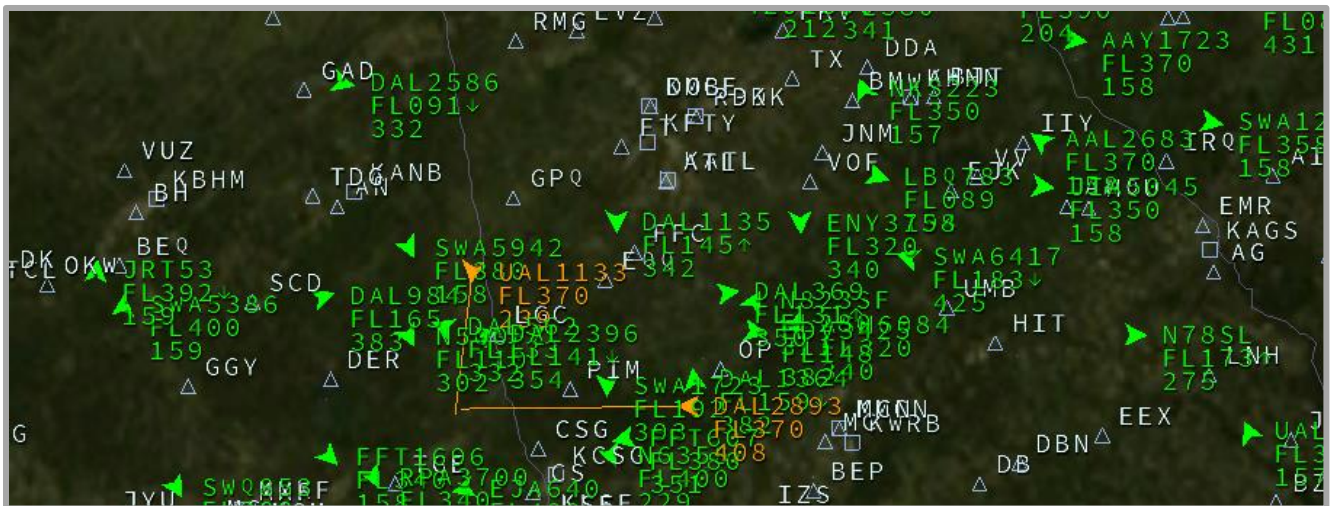

Figure 1. BlueSky simulation of SWIM air traffic scenario around ATL. A two-aircraft conflict is identified (orange) and resolved by an SSD-based CD&R strategy.

$10 \leftrightarrow 28$. According to the *Airplane Flying Handbook* (FAA, 2016), the landing direction is predominantly from the east (using 26R and 27L) so that arriving aircraft can use headwinds for deceleration. Conversely, the takeoff direction is predominantly to the west (using 26L and 27R) such that the incoming and outgoing air traffic flows are aligned in the same direction. Figure 3 shows an an example of how the runway information for ATL (ICAO Airport Code: KATL) is encoded in the BlueSky scenario file after creation of the flights DAL2396 and DAL369 using the CRE command. In addition, the flight parameters heading (HDG), altitude (ALT), and speed (SPD) of the aircraft are updated by aircraft ID when a new SWIM message about that aircraft arrives as shown for flight ENY3758 in Figure 3. Note that the updated positions of each aircraft are simulated using BlueSky rather than being informed from SWIM.
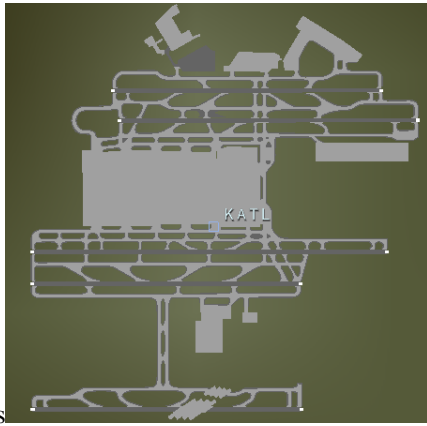

Figure 2. Layout of five runways at ATL airport

The data from 2200 to 2310 UTC, a time considered to be of average air traffic density for ATL, were simulated in BlueSky. For simplicity in this investigation, only the 14 aircraft added to the airspace between 2300 and 2310 UTC were probabilistically interrogated using NESSUS. Further details about interfacing BlueSky and NESSUS are presented in Section 4.

```
02:09.04>CRE          DAL2396    B752      33.019
02:09.04>DAL2396      ORIG       KRSW
02:09.04>DAL2396      DEST       KATL      RW27L
02:09.03>CRE          DAL369     A320      33.207
02:09.03>DAL369       ORIG       KATL      RW27R
02:09.03>DAL369       DEST       MNMG
02:16.14>ENY3758      HDG        177.965
02:16.14>ENY3758      ALT        32000.0
02:16.14>ENY3758      SPD        395.0
```

Figure 3. Sample of BlueSky scenario file commands

3. FLIGHT PLAN FLEXIBILITY (FPF) METRIC

This section derives the "flight plan flexibility" metric from SSD information produced by BlueSky. This metric is not a direct measure of system safety, but it may be embedded in a system safety metric that considers probabilities of events that would require immediate flight plan deviations.

3.1. State-Space Diagrams (SSDs)

Throughout the simulated air traffic scenario around ATL, aircraft must undergo CD&R in order to avoid loss of separation (LoS) with other aircraft. The CD&R strategy aims to represent NextGen capabilities where conflicts can be resolved automatically using on-board ADS-B technologies without the direct intervention of an air traffic controller. A conflict is detected if a LoS is predicted to occur when extrapolating the current aircraft states over a specified look-ahead time. With a look-ahead time set to 5 minutes for tactical maneuvering, Figure 1 shows that conflicts between aircraft (highlighted in orange) are detected. Following conflict detection, a resolution algorithm adapts each aircraft's flight plan to prevent a LoS. There are different rules for prioritizing the flight path adjustments an aircraft may implement. Here, the priority is to first adjust the speed (no more than +/- 10 knots), followed by the heading, and then the altitude. In addition, the CD&R strategies in BlueSky are optimized simultaneously for all aircraft in conflict, making sure to not create additional conflicts with neighboring aircraft (Balasooriyan, 2017). In this air traffic scenario, conflicts were detected and successfully resolved; a LoS did not occur during the simulation.

FPF is computed from the SSD information used for CD&R and measures of air traffic complexity (Balasooriyan, 2017). State-space diagrams are derived from Velocity Obstacles (VOs) which are defined as a velocity (i.e., speed and direction) that will result in a conflict over the look-ahead time. The summation of all the VOs relative to a particular aircraft are shown in Figure 4a and comprise its set of Forbidden Velocities (FVs). In order to maneuver out of conflict, the velocity vector (red arrow) must terminate outside of the set of FVs. In contrast, the performance limits of the aircraft (e.g., minimum and maximum velocities) bound the set of Reachable Velocities (RVs) shown in Figure 4. The intersection of the FVs and RVs is termed as the set of Forbidden Reachable Velocities (FRVs), which are shown alongside their complement - the Allowable Reachable Velocities (ARVs) - in Figure 5.
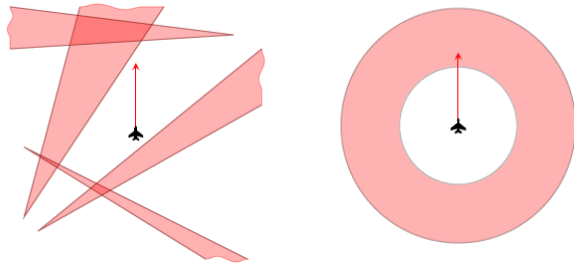
Figure 4. Sets of Forbidden Velocities (FVs) (left) and Reachable Velocities (RVs) (right) relative to a single aircraft (Balasooriyan, 2017)
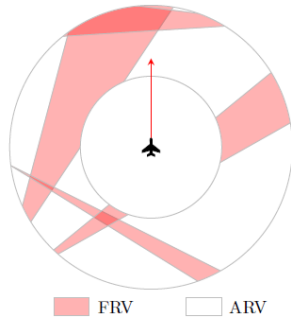

Figure 5. Intersection of FVs and RVs compose Forbidden and Allowable Reachable Velocities (FRVs and ARVs) (Balasooriyan, 2017)

3.2. Computing FPF

The formulated measure of FPF is the ratio of the area of FRVs to the total area of the RVs in Figure 5, which are tracked for each aircraft in a given air traffic scenario. Shown in Eq. (1), this metric results in a flexibility measure ranging from 0 to 1. Thus, FPF represents the flexibility of an individual aircraft to adapt and maneuver through changing flight conditions (due to weather, wind, nearby aircraft, etc.) to avoid LoS.

$$FPF = 1 - \frac{Area(FRV)}{Area(RV)} = 1 - \frac{Area(FRV)}{Area(FRV) + Area(ARV)} \qquad (1)$$

An FPF close to 0 therefore indicates there are only a small fraction of velocities among the RV the aircraft may assume that will not result in a conflict. Conversely, an FPF of 1 means that the aircraft may opt for any velocity within its set of RV and not create any conflicts. An FPF of exactly 0 means that conflict (and eventual LoS) is inevitable if no remedial action is taken by other aircraft in the system (from ATC intervention or a CD&R strategy).

The FPF metric developed in this paper is a new metric, but it does relate to separation distance between aircraft, which is a metric of interest to the FAA. The FPF metric relates to separation distance in that the Forbidden Reachable Velocities are zones in the velocity space that will result in a loss of separation over a standard lookahead time of five minutes. The conflict-resolved flight plan from CD&R and thus the resulting FPF of each aircraft are dependent on the aircraft state variables (position, speed, heading, and altitude). The next task discusses the effect of uncertainty in the initial position of the aircraft on the computed FPF.

4. PROPAGATING UNCERTAINTY THROUGH THE NAS USING NESSUS

The effect of uncertainty in the reported aircraft position on FPF is investigated using the SwRI-developed NESSUS software.

4.1. Uncertainty in Reported Aircraft Position

In this analysis, realistic distributions for aircraft position reported from ADS-B signals were determined from (Jones, 2009), which align with the FAA Advisory Circular No. 20-165 (2010). The standards for positional accuracy are summarized according to Navigational Accuracy Categories for position (NACp) and likewise for velocity as NACv. These categories range from 0 to 10 according to the integrity of the navigational information with each category corresponding to standard deviations of a Gaussian distribution with means at the reported latitude and longitude. For example, an NACp value of 7 corresponds to a maximum position standard deviation of 0.1 NM. In this study, we used standard deviations of latitude and longitude according to NACp values of 4, 5, and 7 as shown in Table 1, however, note that an NACp value of 8 is the typical standard for the Class B airspace around ATL.

Table 1. NACp values and corresponding position standard deviation in NM and degrees

NACp Value	Standard Deviation (NM)	Standard Deviation (degrees)
4	1.0	0.0016
5	0.5	0.008
7	0.1	0.016

4.2. Uncertainty Propagation with NESSUS

NESSUS is used to investigate the impact of uncertainty from ADS-B signals on flight plan flexibility. The general NESSUS analysis configuration wraps around the BlueSky code which is defined as a function of 28 random variables which are the initial latitude and longitude positions of each of the 14 aircraft. Each random variable is modeled probabilistically as a Gaussian (i.e., normal) distribution with the mean defined from the latitude or longitude value reported by SWIM data and a standard deviation (in degrees) according to the chosen NACp value from Table 1. To conduct the analysis at different NACp values the standard

deviations of the random variables were simply modified in the NESSUS GUI.

Two key decisions in an uncertainty propagation analysis are the number of samples to propagate through the simulation and the method to generate the samples. For this uncertainty propagation analysis, 1000 samples of the 28 random variables were generated from a Latin-Hypercube Sampling strategy (McKay, Beckman and Conover, 1979). The choice in the number of samples to generate from a sample design strategy depends primarily on two factors – the computational cost of the simulation and the output statistics desired. In this case, the goal is to capture the mean of the FPF for each aircraft and also have some information on uncertainty bounds for comparison between aircraft and NACp categories. Because each analysis takes approximately 15 seconds, the uncertainty propagation analysis with 1000 samples took approximately 4 hours. This number of samples was deemed sufficient to accurately capture both the mean of the FPFs and their 80% confidence bounds, however the 95% confidence intervals are reported to provide clearer visualization of the uncertainty in the FPFs for each aircraft.

5. RESULTS

Figure 6 shows the scatter in the flight plan flexibility due to the uncertainty in the initial position of the 14 aircraft added to the airspace between 2300 and 2310 UTC. Note that since the airspace around ATL is dense, there is no flight plan among these with complete flexibility (i.e., FPF=1) at any time. Furthermore, there are some instances where flight plan flexibility is zero, which indicates that other aircraft were required to adjust their flight plans to prevent a LoS. There were no LoS occurrences throughout this 10-minute simulation. Further, Figure 6 demonstrates that the flight plan flexibility displays some alternative behaviors in the higher uncertainty cases (i.e., DAL2396 and DAL562) when the CD&R algorithm computed different optimal flight paths for some realizations of the input parameters.

From the results presented in Figure 6, the mean and standard deviation of flexibility is computed at each time instant for each of the three standard deviation cases. The 95% confidence bounds (+/- 2σ) are shown in Figure 7 for six of the fourteen flights in Figure 6. Note that the Gaussian assumption does not hold when multiple flight paths are chosen by the CD&R algorithm. For instance, the FPF results for flight DAL562 in Figure 6 indicate that when σ=0.5 NM (shown in red) and σ=1.0 NM (shown in green), some portion of the FPF responses deviate from the typical FPF. This is
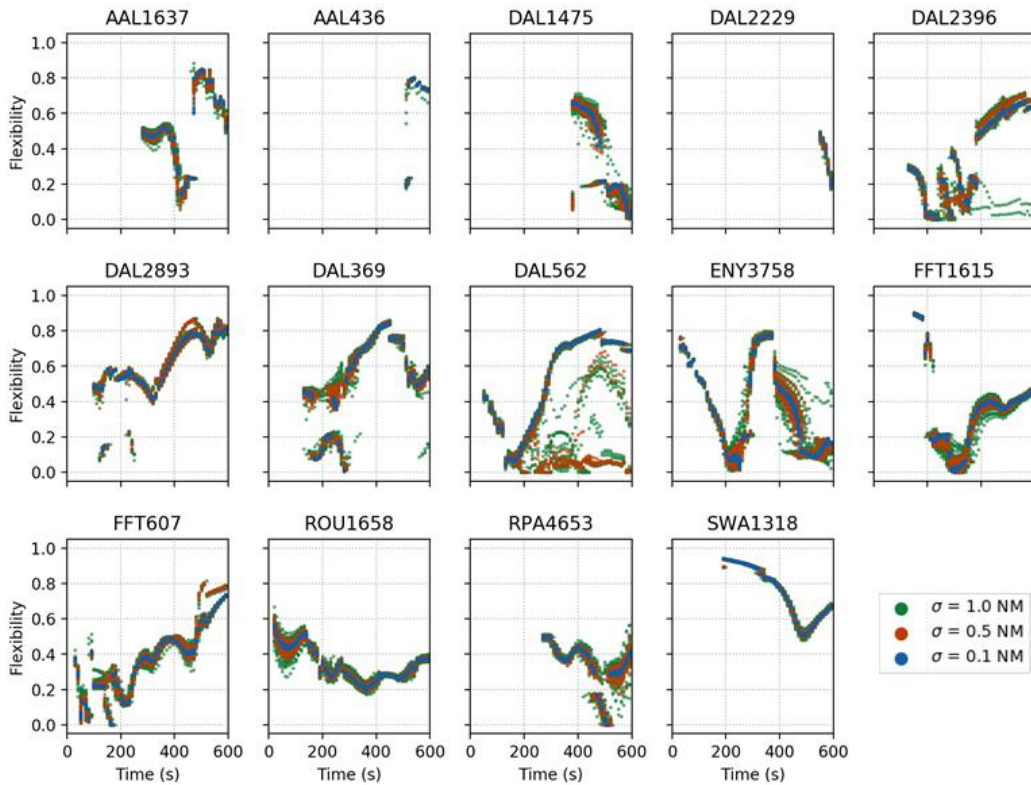

Figure 6. Uncertainty in predicted flight plan flexibilities due to three position uncertainty levels: 0.1 NM (blue), 0.5 NM (red), and 1.0 NM (green) - for 14 aircraft throughout a 10-minute BlueSky air traffic simulation
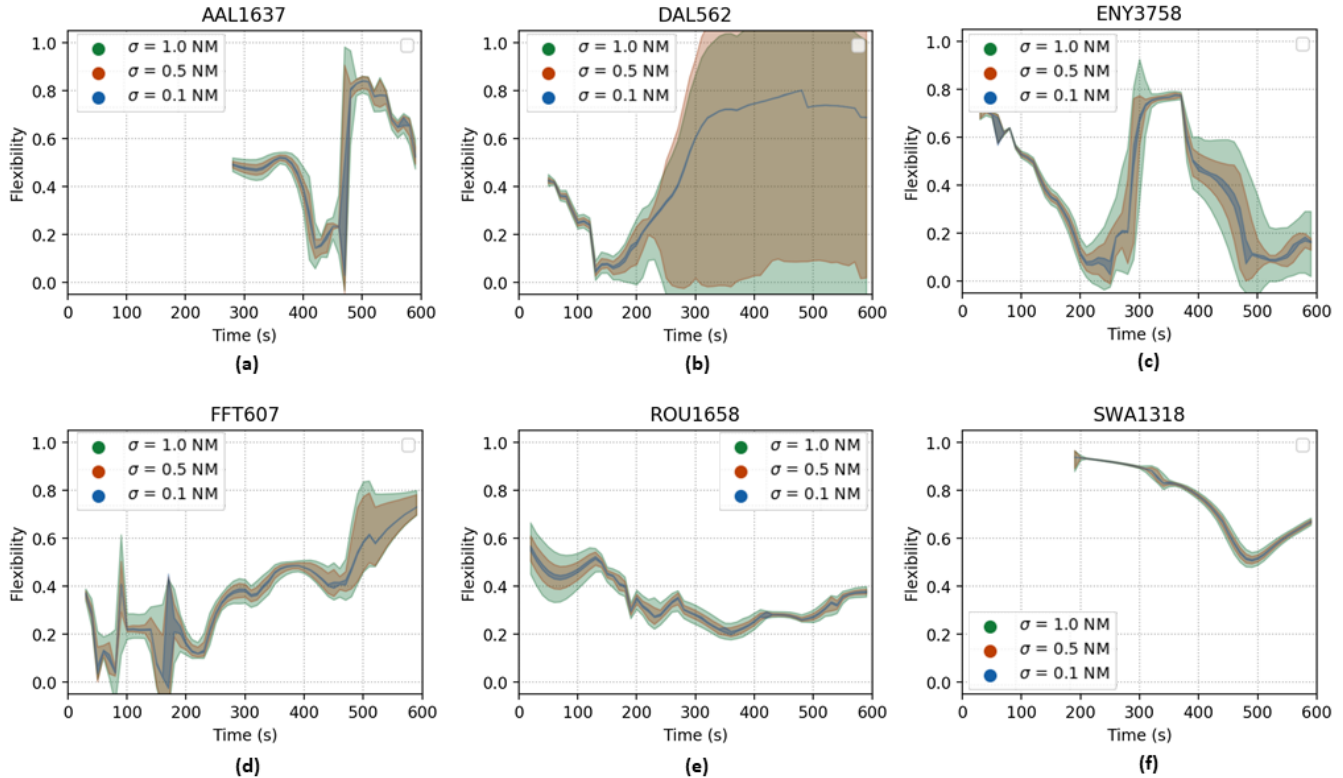
Figure 7. 95% confidence bounds for the predicted aircraft flexibilities of six aircraft (a-f) due to three signal uncertainty levels - 0.1 NM (blue), 0.5 NM (red), and 1.0 NM (green) - throughout a 10-minute BlueSky air traffic simulation

because those samples follow a different flight trajectory than that chosen by CD&R. Figure 7b demonstrates that the 95% confidence bounds generated from the computed mean and standard deviation of the FPF responses of DAL562 in these cases are well outside the FPF bounds of 0 and 1. This is because the standard deviations for these cases are very high and the Gaussian distribution is not constrained by the actual FPF bounds. A more advanced representation of the FPF distribution might be bi-modal, mixed Gaussian, or beta distribution to enforce the FPF bounds. This will be addressed in future work.

In parallel to tracking the flexibility of these fourteen aircraft, the flexibilities of the other 105 aircraft in the airspace were also logged over the 10-minute simulation. These could be used to determine inflexible areas of the airspace that are more susceptible to LoS, which may then inform risk maps of the NAS. In addition, the FPF metric and its uncertainty may be used in a CD&R strategy itself. For example, the goal of the strategy could be for all aircraft to maintain a threshold flexibility (say, 0.1) with a probability of 95%.

6. FUTURE WORK

Of immediate interest is transitioning from BlueSky to the comprehensive gate-to-gate simulation software being developed by Optimal Synthesis, Inc. for prognostics. This simulation software is being expanded as part of this same NASA ULI-funded project, and one of the project goals is to create modules for safety and quantified uncertainties to identify risk and potential mitigation strategies. To create such modules, continued investigation into quantifying sources of uncertainty and NAS safety metrics is planned.

Further research includes aggregating FPF of individual aircraft to form an overall indicator of NAS safety. This is a significant challenge because there are probabilistic correlations and dependencies among the calculated flexibilities that prohibit simple aggregation methods. The goal is to form an overall measure of complexity that has tie-ins to ATC workload and pilot performance – quantities of interest to other research within this project - and may become a component in a composite air traffic safety metric. A composite safety metric might include the likelihood of a flight plan deviating event, number of in-conflict aircraft, and the occurrence of LoS. In addition, there are other safety metrics of interest to the FAA that may be incorporated, one being the Close Approach Probability used by Jones (2009) which considers when ADS-B is in fault-free and fault conditions.

Continued investigations will be conducted to construct realistic probabilistic models of epistemic and aleatory sources of uncertainty in the NAS from available data. The

sources of uncertainty to be explored include weather effects, aircraft performance uncertainty from BADA, and signal and position measurement errors from communication systems in addition to ADS-B. The distinction between aleatory and epistemic sources of uncertainty is that aleatory sources are those that stem from natural variability and are irreducible whereas epistemic sources stem from "lack of knowledge" and their uncertainty can be reduced. This work will feed into other aspects of the ULI by establishing a prior uncertainty on system safety with and related work demonstrating the use of Bayesian Entropy Networks to fuse data from multiple information streams. The purpose of the information fusion is to reduce the epistemic uncertainty in a safety relevant system output with the overarching challenge being providing a continuous update to system safety as data is being collected. Crucial to these Bayesian updating procedures will be the employment of a global sensitivity analysis that will identify the most significant contributors to uncertainty in system safety. Those sources of uncertainty that are most influential will then be candidates for uncertainty reduction during the information fusion and Bayesian updating process.

REFERENCES

Balasooriyan, S. (2017) *Multi-aircraft Conflict Resolution using Velocity Obstacles*. Delft University of Technology.

Federal Aviation Administration (FAA) (2010) *Airworthiness Approval of Automatic Dependent Surveillance - Broadcast (ADS-B) Out Systems*. AC 20-165.

Federal Aviation Administration (FAA) (2016) *Airplane Flying Handbook*. FAA-H-8083-3B. Oklahoma City, OK: United States Department of Transportation.

Hoekstra, J. M. and Ellerbroek, J. (2016) 'BlueSky ATC Simulator Project: An Open Data and Open Source Approach', in *Proceedings of 7th International Conference on Research in Air Transportation*. Philadelphia, PA.

Jones, S. R. (2009) *ADS-B Surveillance Separation Error Sensitivity Analysis*. 2nd edn.

McKay, M. D., Beckman, R. J. and Conover, W. J. (1979) 'A comparison of three methods for selecting values of input variables in the analysis of output from a computer code', *Technometrics*, 21(2), pp. 239–245.

BIOGRAPHIES

Erin C. DeCarlo is a Research Engineer in the Computational Materials Integrity section within the Materials Engineering department at SwRI. She obtained her bachelor's degree in Civil Engineering from Tennessee Technological University and her master's and Ph.D. from Vanderbilt University under the direction of Sankaran Mahadevan. Her research at SwRI focuses on applying probabilistic methods and Bayesian techniques to conduct reliability analyses of complex material and mechanical systems.

Barron J. Bichon manages the Computational Materials Integrity team within the Materials Engineering department at SwRI. He obtained his B.S. in from The University of Memphis, his M.S. from the University of Illinois at Urbana-Champaign, and his Ph.D. from Vanderbilt University (all in Civil Engineering). His work focuses on developing and applying methods to enable confident decision-making using computational models and experimental data.